IT Unit 3

Topic 2

# Ch 3, Data Analytics: Drawing Conclusions, Part 2

**Identifying patterns and relationships between data**, 152-165

1. What is the purpose of using statistics when working with large data sets?

* To convert large quantities of raw data into small, informative, meaningful summaries, and allow patterns & relationships to be identified.

1. List the 3 most useful statistical concepts from table 3.7, p 152.

* Average
* Standard deviation, which indicates how consistent the values are in a data set
* Correlation, which show the trend of data sets

**Basic statistics**

1. Distinguish between the three methods of calculating an average.

* The mean, which is the average figure across the whole data set.
* The median, which is the middle number when the data is placed in order.
* The mode, which is the value that occurs most frequently.

1. What is the purpose of the standard deviation?

* To show how consistent the values are compared to the mean in a data set and also to see if there are seasonal trends.

**Correlation and causality**, p 156

1. Why do you need to be careful when looking at patterns in data in terms of cause and effect?

* Because you might misinterpret what the data actually means and you may also think that it’s a causal relationship when really it’s not. You can only tell if it’s a causal relationship if you conduct further research.

**Data visualisations, p 157**

1. What are data visualisations used for?

* To make data easier to interpret, explain, understand trends and patterns and also analyse.

**Queries and searches,** p 159

1. How can queries, searches, filtering and sorting be used when analysing large data sets?

* They can be used to help you locate patterns in your data sets so that you can find the larger story that your Outcome data is telling you.

**Conditional formatting,** p 164

1. Explain the process of conditional formatting.

* Conditional formatting allows you to change the appearance of data automatically based on its current value in both spreadsheets and databases. This highlights data of special interests and can also point out errors in the data.